Bayesian Bivariate Markov-Switching Multifractal model
Examining the inter-frequency dependence between crude oil spot and futures volatility

Jan Tomoya Greve
https://github.com/Jantg
Department of Statistical Science,
Duke University

Jayeon Lee
https://github.com/jaeyeonfr
Department of Statistical Science,
Duke University

1 Motivation

Markov-Switching Multifractal model (MSM) introduced in Calvet and Fisher (2001[2], 2004[3]) is a reinterpretation of Multifractal Model of Asset Returns (MMAR) by Mandelbrot (1997)[1] in stochastic volatility model setting. Because of this construction, it shares the same characteristics enjoyed by MMAR such as scaling, fat tails, volatility clustering and long memory which are well known stylized facts of financial data. Moreover, it circumvents the combinatorial difficulty of MMAR and ensures a strict stationarity in the process, thus making it more suitable for econometric settings. Finally, by construction it induces a decomposition of volatility into several Markovian states with varying degree of switching frequencies (from high to low in ordered manner) which has a considerable inferential appeal.

However, its bivariate extension by Calvet and Fisher (2006)[4] albeit parsimonious assumes that a single parameter governs the dependence of switchings between two series across all frequencies simultaneously, thus diminishing its inferential capabilities enjoyed in the univariate case. Economically, robustness against shocks will depend on the nature of stocks and therefore, changes in volatility of stocks with similar fundamentals may occur in different frequencies thus making it harder to capture with their model. Introduction of Coupl Optional Pólya Tree prior in joint estimation phase (which will be detailed in later chapter) is exactly meant to tackle this problem and thereby achieving more breadth in inference to uncover possible linkage between series in some frequencies and aid predictions.

2 Univariate MSM by Calvet and Fisher (2004)[3]

Here, we introduce the discrete univariate MSM (from here on when we mention MSM it will always be discrete MSM). Univariate MSM requires only 4 parameters, thus compared to other stochastic volatility models has a parsimonious specification.

First, we introduce the multipliers which are first-order Markov state with $\bar{k}$ components

$$M_t = (M_{1,t}, ..., M_{\bar{k},t}) \in \mathbb{R}_+^{\bar{k}}.$$

Each components have the same marginal distribution $M(\theta)$, but the switching frequency is given by exogenous parameters $\gamma \equiv (\gamma_1, \gamma_2, ..., \gamma_{\bar{k}})$. Therefore the dynamics of multipliers are specified as following

$$M_{k,t} = \begin{cases} m \sim M(\theta) & \text{with probability } \gamma_k \\ M_{k,t-1} & \text{with probability } 1-\gamma_k \end{cases}$$

where the switching probability γ_k is specified as

$$\gamma_k = 1 - (1-\gamma_1)^{b^{k-1}}.$$

Due to this construction, multipliers are mutually independent with each other. Then the stochastic volatiliy is specified as follows

$$\sigma(M_t) \equiv \bar{\sigma}\Big(\prod_{i=1}^{\bar{k}} M_{k,t}\Big)^{\frac{1}{2}},$$

where $\bar{\sigma}$ is a positive constant to be estimated. Finally, returns r_t are

$$r_t = \sigma(M_t)\varepsilon_t \quad \text{where } \varepsilon_t \sim N(0,1).$$

The full parameter vector is $\psi \equiv (\theta, \bar{\sigma}, b, \gamma_1)$ which exemplifies the parsimony of the model. Resulting graph from $t-1$ to t are presented above.

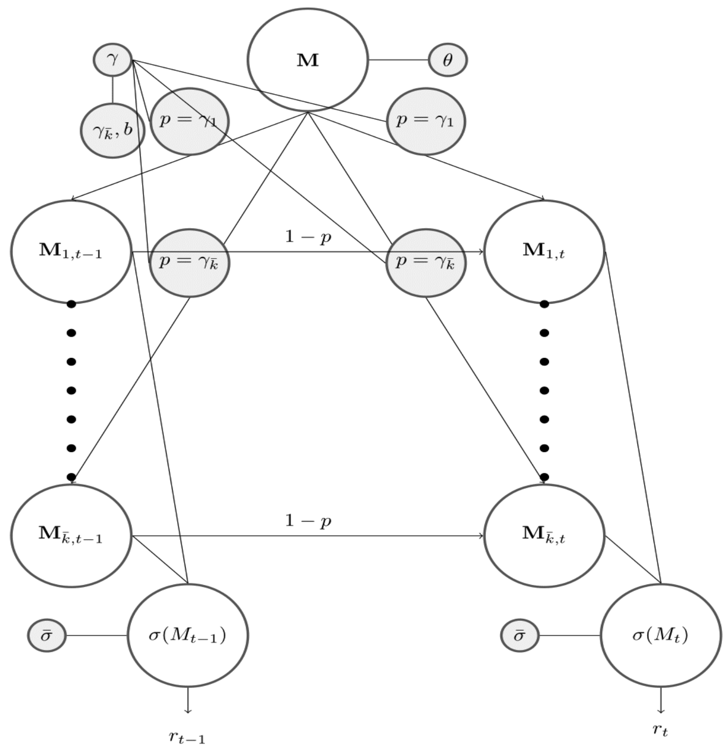

Fig1: Univariate MSM($\bar{k}$).Grey circles represent deterministic nodes.

For the distribution of M, the only restriction is that the values must be positive and that $\mathbb{E}(M) = 1$. This completes the specification of the univariate MSM and the model is denoted as MSM($\bar{k}$) where the choice of $\bar{k}$ is a model selection problem rather than the estimation.

3 Bivariate MSM by Calvet and Fisher (2006)[4]

In Calvet and Fisher (2006)[4], the bivariate extension of MSM($\bar{k}$) for series α, β is done by the full parameter vector $\psi \equiv (\theta^\alpha, \theta^\beta, \bar{\sigma}^\alpha, \bar{\sigma}^\beta, b, \gamma_1, \rho_\varepsilon, \lambda)$ where ρ_ε governs the correlation between observation error ε^α and ε^β while λ controls the correlation of switchings between multipliers in two series with the same frequency. In more detail, we define the arrival of new state for multiplier k in series α, β as $1^\alpha_{k,t}, 1^\beta_{k,t} \sim$ Bernoulli(γ_k). State transition will then be $M_{k,t} \sim M((\theta^\alpha, \theta^\beta))$ if both indicators are 1 and if only one $c \in \{\alpha, \beta\}$, then $M^c_{k,t}$ will drawn from the marginal M^c and else will stay the same for two. Then, we set $\Pr(1^\alpha_{k,t} = 1) = \gamma_k$ and $\Pr(1^\beta_{k,t} = 1 | 1^\alpha_{k,t} = 1) = (1-\lambda)\gamma_k + \lambda$. By similar procedure, the distribution of arrivals between these two series in k can be completely specified.Then, λ will be estimated across all frequencies of $k = 1, ..., \bar{k}$.

Finally, we introduce the two step estimation procedure of all these parameters. In stage 1 parameters $(\theta^\alpha, \theta^\beta, \bar{\sigma}^\alpha, \bar{\sigma}^\beta, b, \gamma_1)$ will be estimated by maximizing the sum of univariate likelihoods. Then in stage 2, using the estimate of these parameters (let's denote those as ψ_1), simulated bivariate likelihood from particle filters $L(r^\alpha_t, r^\beta_t, \psi_1, \rho_\varepsilon, \lambda)$ will be maximized.

4 Bayesian Bivariate MSM using Coupled Optional Pólya Tree prior

4.1 Brief description of the proposal

Here, we introduce the Bayesian Bivariate MSM which builds upon Calvet and Fisher (2006) and improves one crucial inferential aspect of it while also providing some improvements in the forecasting as well.

In the original approach, it was assumed that switching in one frequency of the multiplier in α can only be correlated with the one in β with the same frequency. However, economically we might expect that regime switching in the volatility component can have different degree of magnitude depending on the nature of the stock but still be correlated, which is hard to capture in the original setting of bivariate MSM. For example, spot price of a commodity and its futures are very likely to have high correlation in the volatility, but the switchings of the multipliers in spot price may only occur in higher frequency components while that of futures could occur in much broader frequencies due to its nature of being more sensitive to the change in expectations.

To capture such heterogeneity in switchings, we would rather utilize the independence of switchings within the same series and construct $2^{\bar{k}} * 2^{\bar{k}}$ contingency table of switchings and give it a prior so that after the first stage estimation of parameters for two decoupled univariate series are done with particle filters, we can sample from the estimated parameters to update the joint prior of volatility arrivals. The contingency table where we plan to give a prior would thus have the following structure For this task, we

	$1^{\beta}_{1,t}=0,...,1^{\beta}_{\bar{k},t}=0$	$1^{\beta}_{1,t}=0,...,1^{\beta}_{\bar{k},t}=1$	$\cdots$
$1^{\alpha}_{1,t}=0,...,1^{\alpha}_{\bar{k},t}=0$	$\cdot$	$\cdot$	$\cdots$
$1^{\alpha}_{1,t}=0,...,1^{\alpha}_{\bar{k},t}=1$	$\cdot$	$\cdot$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$

propose to use Coupling optional Pólya tree prior (from here on we call it co-OPT) due to its several advantages over other priors which we will briefly explain. One of the major advantage of this prior is that not only does it provide measure of closeness of two samples, it can also learn the structure of the source of heterogeneity by for example identifying factors that contributed to the similarity. This suits well to our objective which is not just obtaining a single measure of dependence between switchings in two series but rather learn the whole structure of it for inference and for predictions. Additionally, according to Ma and Wong (2011)[5], this prior has a considerable computational advantage over priors which handles similar two-sample problems such as dependent Dirichelet prior.

In fact, this specification closely parallels the example 5 in Ma and Wong (2011)[5] with sampled data on a $2^{15} * 2$ contingency table with 15 iid predictors $X_1,...,X_{15} \sim \text{Bernoulli}(0.5)$ and a binary response Y which is a mixture of Bernoulli distributions depending on realizations of some X_i's.

4.2 Thoughts on practical implementation of our procedure

The implementation of this model is done as follows.

1. Decoupled estimation of series level parameters:
   For each series $i=\alpha,\beta$ we estimate $\bar{\sigma}^i, \theta^i, \gamma^i_1, b^i$ with particle filter.

2. Recoupled joint esitimation:
   After obtaining parameter estimates in step 1, we sample from each series and record the arrivals $1^i_{k,t}$ $\forall i,k$ in the contingency table above and update the co-OPT prior by giving one series the role of a predictor (with $\bar{k}$ features) and the other as a response (which is categorical with $2^{\bar{k}}$ values). Additionally, the other parameter ρ_ε can be obtained from the observation error $(\varepsilon^{\alpha}, \varepsilon^{\beta})$ covariance matrix which we can give Inverse-Wishart prior to.

3. Inference:
   With sufficient amount of data at each time t, the co-OPT prior will capture the underlying structure of the difference. Then, as in Ma and Wong (2011)[5], it can be recovered by the hMAP tree topology for inferring the conditional relationship of one series to the other.

4. Forecasting:
   For the predictor series, the forecasting will be done solely by its series specific parameters. For the response series however, we can take samples of partition trees from the posterior of co-OPT and draw predictive samples conditioned on the samples of the predictor series (if there are any interactive structure among some predictors and the response in the sampled tree).

5 Summary of our proposals and possible additional work in consideration

In this proposal, we introduced the univariate and bivariate MSM and highlighted a possible improvement in the bivariate case by utilizing the Coupling Optional Pólya tree prior in the second stage of the estimation (the recoupled joint estimation part) and thereby rephrasing the model into Bayesian setting. It will enable us to infer the structure of dependence among the multipliers with different frequencies for two series, thereby taking into account the heterogeneity in the magnitude of the regime switch between series.

We would like to construct this model and apply it to crude oil spot and futures volatility which often has a multifractal structure (as seen the in the figure below with the comparison with the brownian noise) and is of great economic importance. Indeed, MSM has already been applied to crude oil returns such as Wang et al (2016)[6] and there is a demand for more thorough bivariate analysis.

Finally, considering the nature of MSM which closely follows various stylized facts of financial data in a consistent manner (consistent meaning without introducing ad-hoc features) it does have a natural appeal to be utilized in predictions of the future volatility. Therefore, conducting a portfolio optimization using the developed Bayesian Bivariate MSM is of our interest.

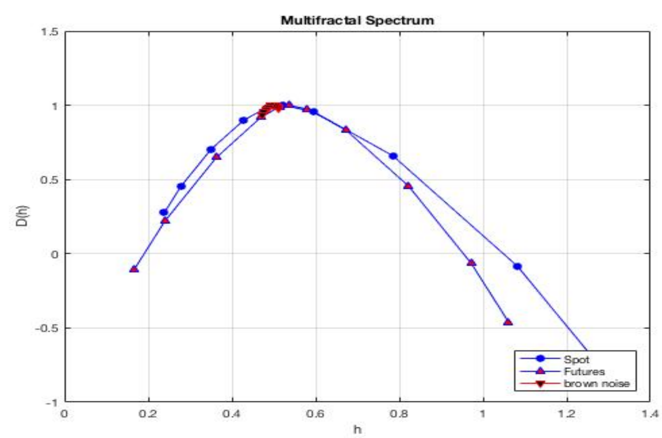

Fig2: Multifractal Analysis of crude oil spot and future returns from May 1983 to date. Wide support suggest that those are multifractals

[1] A. Fisher B. Mandelbrot and L. Calvet. A multifractal model of asset returns. *Cowles Foundation Discussion Paper*, (1164), 1997.

[2] L. Calvet and A.Fisher. Forecasting multifractal volatility. *Journal of econometrics*, 105(1):27–58, 2001.

[3] L. Calvet and A.Fisher. How to forecast long-run volatility: regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2(1):49–83, 2004.

[4] A.Fisher L. Calvet and S.B. Thompson. Volatility comovement: a multifrequency approach. *Journal of econometrics*, 131(1).

[5] L. MA and W.H. Wong. Coupling optional pÃşlya trees and the two sample problem. *Journal of the American Statistical Association*, 106 (496):1553–1565, 2011.

[6] WU C. Wang, Y. and L. Yang. Forecasting crude oil market volatility: A markov switching multifractal volatility approach. *International Journal of Forecasting*, 32(1):1–9, 2016.